R842-05 ERIC REPORT RESUME

ED 010 243 3-01-67 24 (REV)

ANALYSIS OF WORD FREQUENCIES IN SPOKEN LANGUAGE OF CHILDREN.

BEIER, ERNST G. * AND OTHERS
WXS77656 UNIVERSITY OF UTAH, SALT LAKE CITY

CRP-S-277
BR-5-8036
- -65
EDRS PRICE MF-$0.09 HC-$0.84 21P.

LANGUAGE RECORDS (PHONOGRAPHY), *LANGUAGES, LANGUAGE FLUENCY, *LANGUAGE RESEARCH, LANGUAGE ABILITY, *VOCABULARY DEVELOPMENT, *CHILDREN, TAPE RECORDERS, SALT LAKE CITY, UTAH

AN ATTEMPT WAS MADE TO ESTABLISH CERTAIN BASE RATES IN LANGUAGE USAGE OF CHILDREN AND TO INVESTIGATE THEIR PSYCHOLOGICAL SIGNIFICANCE. SPECIFICALLY THE STUDY WAS AIMED AT DISCOVERING WHETHER THE FINDINGS OF G.K. ZIPF (1965) HELD FOR SPOKEN LANGUAGE OF CHILDREN AND WHETHER AGE DIFFERENCES INFLUENCED THE RELATIONSHIP OF VARIETY AND FREQUENCY OF WORD OCCURRENCES. ZIPF STATED THAT A RELATIONSHIP EXISTS BETWEEN HIGH FREQUENCY, SMALL VARIETY, AND SHORTNESS OF WORDS, A RELATIONSHIP PRESUMABLY VALID FOR LANGUAGE IN GENERAL. IN THIS STUDY 12- AND 16-YEAR-OLD BOYS (N 15 FOR EACH AGE GROUP) OF AVERAGE INTELLIGENCE WERE GIVEN A TAPE RECORDER TO OBTAIN THEIR VERBAL OUTPUT. FORTY THOUSAND WORDS OF EACH OF THE GROUPS WERE ANALYZED WITH THE STARKWEATHER PROGRAM ON AN IBM 7094 COMPUTER. THE RESULTS WERE PRESENTED IN TERMS OF LISTS OF WORDS USED, RATIOS OF THE NUMBER OF DIFFERENT WORDS SPOKEN TO THE NUMBER OF TOTAL WORDS, AND THE RATIO OF VARIETY TO FREQUENCY OF OCCURRENCE. (JC)

ED010243

5-8036

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Analysis of Word Frequencies in Spoken Language of Children[1, 2]

Ernst G. Beier - University of Utah

John A. Starkweather - University of California

Don E. Miller - University of Utah

Abstract

Zipf (1965) states that a statistical relationship has been established between high frequency, small variety and shortness in length of words, a relationship which is presumably valid for language in general. Zipf based his work on the analysis of written language. The present study is concerned with discovering whether this law also holds for spoken language of children and if age differences influence this relationship of variety and frequency of occurrence of words, as well as the frequency of specific word groups (such as negative words, self-reference etc.). For this purpose 15 12-year old boys and 15 16-year old boys of average intelligence were given a small tape recorder to obtain their verbal output. Forty thousand words of each of the groups were analyzed with the Starkweather program on an IBM 7094 computer. The results are presented in terms of lists of words used, the ratios of the number of different words spoken to the number of total words, the ratio of variety to frequency of occurrence. It was thought that this study contributes to a better understanding of children's spoken language and the growth of their available vocabulary.

---

[1] This research was supported by the U.S. Office of Education, Project No. S-277.

[2] We wish to express our appreciation to Mr. Arthur E. Arnesen, Administrative Assistant to the Superintendent of Schools, Salt Lake City for his cooperation in collecting this data.

Analysis of Word Frequencies in Spoken Language of Children

A Model for an Intercultural Study

Ernst G. Beier - University of Utah

John A. Starkweather - University of California

Don E. Miller - University of Utah

Introduction

This study was concerned with establishing certain base rates in language usage of children and investigating some of their psychological significance.

Related Literature

A number of studies have investigated language behavior in children with reference to specific variables such as diversity, socio-economic background, the type-token ratio (TTR) and others. Chotlos (1944) investigated the effects of age, socio-economic level, IQ and sex on language behavior. He computed the TTR from 108 Ss selected from 1000 children between eight and 18 years of age and found that diversity as measured with the TTR increased with age and IQ, but seemed independent of location and sex. Busemann's (1925) "action quotient" (i.e. verbs as related to adjectives, nouns and participles) also increased with age and IQ. Bernstein (1962) reported a language study carefully controlled for social class differences. He postulated different verbal planning, orientation in different social classes and proposed an "elaborated" and "restricted" linguistic code which is based on phrase length, word length and pausing. He found social class differences, with lower socio-economic class members using longer mean phrase length, less pausing, but (with IQ equated) no shorter word length. Pringle and Tanner (1958)

recorded conversation of pre-school children in controlled and spontaneous conditions. They found deprivation in language skill development of pre-school children in care homes. This conclusion was also supported by McCarthy (1954), Verplank (1955) and Milner (1951). Minifie, Darley and Sherman (1963) obtained three different language samples from five and eight-year old children using picture cards as stimuli. They found relatively low temporal reliability at both age levels, which perhaps is due to the small sampling of language. Smith (1926), who recorded spontaneous conversation of 88 children, noted a great variability within age groups and suggested that other factors than age would be responsible for the variance. Lorge-Thorndike word frequency scales (1944) were based on written language, and so is Fraprie's (1950) scale of most common words. West (1953) and Rosenzweig and McNeill (1962) pointed out important differences between the scales. Zipf's work (1935) which gave us Zipf's laws support that frequency of occurrence of words is inversely related to word length as well as the total number of words of a given sample. Voelker (1942) reported the 1000 most frequent words extracted from 100,000 words of spoken language of high school seniors, college freshmen and seniors. This work will complement the proposed research.

A number of investigators have looked into the relationship of oral and written language. Horowitz and Newman (1964) and Fraisse and Breyton (1959) found a greater variety of words and specifically more verbs in oral speech. Moscovici and Humbert (1960), however, found that Zipf's laws were supported with his small sample of both oral and written language probes. Strickland (1962) reported that "oral language in children is more

advanced than the language of the books in which they are taught to read." Drieman (1962) took written and oral tests from a small sample of psychology students and found in their written language: longer words, fewer one-syllable words, more different words of one syllable, a more varied vocabulary. Word frequency studies are also used for diagnostic purposes. Gleser, Goldine and Gollschalk (1959) found in five minute word samples with 90 Ss, differences correlated with intelligence and sex. Correlation between word frequencies and intelligence were also found by Zipf (1937). Fairbanks (1944) computed TTR for 30 100-word segments of spoken language from schizophrenics and college freshman. He found the TTR significantly lower for the patients, a study also supported by Mann (1944). Darley, Sherman and Siegel (1959) reported an interesting new look at word frequencies from a diagnostic point of view. They had 35 judges rank 572 frequent nouns, verbs and adjectives in terms of their scale values on a "level of abstraction," and found that these words could be reliably scaled. Flesch (1950) built a more complex reliability score based on a relationship of "definite" to "total" words. Such methods may be useful to utilize word usage of children to assess the tools they have available for learning as well as specific deficits such as may result in dropping out of school.

The new instruments available for measurement of language development such as recording devices and computer programs have been utilized for various important efforts, but apparently no normative data on word frequencies of spoken language such as Lorge-Thorndike prepared for written language are yet available.

Zipf (1965) derived his laws on the basis of written language; we were interested in testing some of his laws on spoken language of children. Since Zipf compares the lawfulness of language on different languages, such as Chinese, Latin, German and English, our model may also serve inter-cultural comparisons. The use of spoken language as raw material became practical, when the necessary instruments became available. These instruments included a small pocket recording instrument which could record 4 hours of speech and a vocal analysis computer program developed by Starkweather (1964) which prints out word frequencies from texts typed on IBM cards.

In our present study, we recorded samples of spoken language of thirty boys, all of normal intelligence (90-110) as measured by 3 subtests of the Wechsler (information, similarity, vocabulary). Fifteen of the boys were 12 years old and in the 6th grade, the other 15 boys were 16 years old and in the tenth grade. We instructed the boys to use the recorders after school, not to give any special speeches or read into the recorder. When we had 5000 words recorded for each boy, we selected about 2700 words and typed them on IBM cards. Altogether we had some 80,000 words for processing. The computer provided us with individual printouts as well as data of several word lists, such as the positive words, (yes, okay) negative words (no, never, none, etc.) self-reference singular (I, me, mine) self-reference plural (we, our) other references (you, they, them) question words (why, who, what, where) the most frequent 1-4 letter words, and the type token ratio (No. of different words over No. of total words). Here we were primarily concerned with the following questions: Are Zipf's findings derived from written language, applicable to spoken language? The two laws

with which we are concerned are (1) the number of different words used increases as the frequencies of occurrence becomes smaller, and (2) that the magnitude of words tends to stand in an inverse relationship to the number of occurrences. In addition, we shall inspect our data to discover just precisely how our two age groups differ on the various variables under investigation.

Results

Zipf (page 24) quotes Eldrige for samples of American newspaper English totaling 43,989 words representing 6002 different words. We shall compare his sample with our sample of spoken language of 6th and 10th grades.

Table 1

Comparison of Frequency of Occurrence with Number of Different Words Used

Freq. of Occurrences	Eldrige Newspaper English Sample Total No. of Wds.: 43,989 No. of Diff. Wds.: 6,002 — No. of Diff. Words	Beier Spoken Language Sample 6th Grade Total No. of Wds.: 42,924 No. of Diff. Wds.: 3,096 — No. of Diff. Words	Beier Spoken Language Sample 10th Grade Total No. of Wds.: 43,406 No. of Diff. Wds.: 3,121 — No. of Diff. Words
1	2976	1314	1374
2	1079	499	522
3	516	291	277
4	294	168	166
5	212	102	107
6	151	102	92
7	105	60	59
8	84	40	41
9	86	44	42
10	45	36	32
11	40	32	31
12	37	23	20
13	25	16	15
14	28	18	26
15	26	16	12
16	17	22	11
17	18	18	22
18	10	14	10
19	15	14	12
20	16	20	13
21	13	8	7
22	11	6	9
23	6	12	4
24	8	6	3
25	6	8	7
26	10	9	11
27	9	2	5
28	6	6	7
29	5	6	8
30	4	7	5

Freq. of Occurrences	Eldrige No. of Diff. Words	Beier 6th Grade No. of Diff. Words	Beier 10th Grade No. of Diff. Words
31	6	3	4
32	4	4	2
33	6	3	1
34	2	4	4
35	5	2	1
36	3	10	4
37	2	4	3
39	2	3	2
40	4	1	2
41	1	3	1
42	7	1	3
43	1	3	2
44	4	2	1
45	1	2	3
46	2	5	2
47	5	4	4
48	1	1	3
49	3	3	3
50	3	1	1
51	1	1	1
52	3	2	1
54	1	2	5
55	1	2	3
56	1	2	1
58	2	3	1
60	1	3	1

Inspecting Table 1 we can see that our samples behave very similarly to the Eldrige sample quoted by Zipf. Clearly, the number of different words increases as the frequency of occurrence decreases. We also note some interesting differences. The spoken language of these children encompasses only about one half of the total different words used. This great paucity of language may be due to age, or it may be due to the fact that we are measuring spoken language. We note with particular interest that the number of different words which occur only once in the spoken language sample is disproportionately smaller than such single words used in the newspaper sample. Apparently our sample did cut down on variety of expression. As the older grade has a slight increase in the number of different words used, we reason that at least part of this variety may be due to age. This does not exclude the possibility that written language as such enhances differenciation expression over spoken language.

The magnitude of words in Eldrige's data appears to rest on an estimate of the word size. His estimate is based on an average number of phonemes (individual sounds). We, however, shall use a somewhat simpler measure (letters) which would not be exactly equivalent but would serve our purpose of presenting an estimate about the magnitude of words in spoken language. In table 2 we shall summarize the Eldrige data and present the data for our sample.

Table 2

Comparison of magnitude of words and frequency of occurrence

Frequency of Occurrence	Eldrige Newspaper Magnitude	Beier 6th grade	Beier 10th grade
1- 4	6 ph	5.4 letters	5.5 letters
5-10	5 "	4.7 "	4.8 "
15-20	4 "	4.1 "	4.6 "
21-30	3.5 "	3.7 "	4.2 "
30-50	3.9 "	3.4 "	4.0 "
51-60	3.3 "	3.3 "	3.9 "
61+	2.7 "	2.5 "	2.7 "

While we used letters rather than phonemes to estimate the magnitude of the words, we obtained again a somewhat similar grading for children's spoken language as compared with newspaper English. A decrease in magnitude as related to an increase of frequency is certainly observable. It is interesting to note that in our approximation we also discovered that age influences the magnitude; the 6th grade has a somewhat broader distribution of shorter words as related to frequency of occurrence. It should be noted, however, that the total number of short words used remains relatively alike in both samples.

Table 3

Rate of short word usage by grade

(Most frequent)	6th grade N	Rate	10th grade N	Rate
1 letter words	2373	0.055	2664	0.061
2 letter words	8544	0.199	8605	0.198
3 letter words	10140	0.236	9975	0.230
4 letter words	6528	0.152	6859	0.158
Total word sample	42924		43406	

It appears that both grades use short words relatively evenly, with a slightly higher use of 1 letter words in the higher grade. We shall now inspect, in answer to our second group of questions, the frequencies of our word lists.

Table 4

A comparison of word lists by grade

Lists	All		6th Grade		10th Grade	
	Words	Rate	Words	Rate	Words	Rate
Positive	1326	0.015	513	0.012	813	0.019
Negative	2442	0.028	931	0.022	1511	0.035
Singular,self	3950	0.046	1712	0.040	2238	0.052
Plural,self	1311	0.015	822	0.019	489	0.011
Others	6716	0.078	2962	0.069	3754	0.086
Question	1311	0.015	522	0.012	789	0.018
Total Words	86,329		42,924		43,406	
Total Different Words	4,567		3,096		3,121	
TTR	0.053		0.072		0.072	

In this table we discover that the tenth grade uses both positive words, negative words, singular self-reference, other reference to other, and question words more frequently than the sixth grade. The sixth grades use only plural self-reference more frequently. We prepared a correlation matrix of these frequencies which is presented in Table 5.

Table 5

Correlation Matrix of Children's Word Frequencies as Related to Grade

Grade	Grade	WPM	1 Syl.	Pos.	Neg.	S.R.S.	S.R.P.	Oth.	Qu.	TT
Words per Minute	59	--	--	--	--	--	--	--	--	--
1 Syllable Words	-15	4	--	--	--	--	--	--	--	--
Positive	41	17	-11	--	--	--	--	--	--	--
Negative	70	44	-14	42	--	--	--	--	--	--
Self-Reference Sing.	36	17	4	15	65	--	--	--	--	--
Self-Reference Plural	-35	-23	42	-33	-52	-46	--	--	--	--
Other	59	30	-30	38	50	34	-67	--	--	--
Question Words	38	0	-29	58	48	44	-40	53	--	--
Type Token	4	0	-39	-9	-5	0	-35	23	3	--

.296 = 05 level of significance
.349 = 025
.409 = 01
.449 = .005

When we interpret the coefficients which are significant at the 0.1 level of significance, we find that the older boys:

1. Speak faster than the younger boys.
2. They use significantly more "positive" words.
3. They use significantly more "negative" words.
4. Slightly more "singular" self-reference.
5. Slightly less "plural" self-reference.
6. They use more "other" references.
7. They use slightly more 'question' words.

It is interesting to note that with intelligence equated, the type token ratio does not differenciate the boys.

We prepared a factor analysis from this data (including some variables not discussed here) which yielded the following factors:

FACTORS
(including variables loading .30 or more)

1. One syllable one letter words .85, 3 letter one syllable words -49, negative 61, self-reference singular 91, self-reference plural -48, question 30, and -59, I 91, we -52, you 36, the -60, that 44.

This factor seems to describe people who are self oriented, use a lot of I or self-reference singular words. They don't refer to "us" or "we" but to "you" they use more negative words. This could be called an egocentric factor. They also ask more questions.

2. Two letter one syllable words 38, total one syllable words 39, plural self-reference 59, other reference -42, words used once -87, type token -89, we 57, information -43.

This factor seems to describe a less bright group, (have less information on hand). They refer more to "we," and "us," but not others. They use less words once, have a lower type token ratio. This could be a "closed group" factor describing close group oriented people who don't have as much access to a variety of language usages.

3. Grade -84, achievement -54, IQ -73, positive -39, negative -53, other reference -46, age -91, you 30, information 34.

This factor seems to describe a less bright, younger, 6th grade group who use less negative and positive words, use fewer references to others, (including "you"). However, they have a higher information score.

4. Positive 67, self-reference plural -32, other reference 56, question 80, total words 48, and -60, you 77, the -33, is 41, A 38, information 38.

This factor seems to describe an "other oriented" group who use more positive words, use more question words, use shorter words, use "you" much more often.

5. WPM 32, 2 letter 1 syllable words -76, 3 letter one syllable words 66, positive -34, the 33, it -32, is -66.

This factor seems to describe a group of somewhat faster talkers who use less positive words, less "is," "it," more "the." Use less 2 letter one syllable words, more 3 letter one syllable words. Not too meaningful a factor.

6. IQ 33, 1 letter single syllable 40, A 87, information 45.

This factor describes a brighter group who use "a" more, use more 1 letter single syllable words.

7. Achievement -64, IQ -46, modified achievement -82, vocabulary -72.

This factor seems to describe a group of low achievers and low IQ (less bright) people.

8. Total one syllable words 76, vocabulary -46.

People who use a lot of one syllable words get lower WAIS vocabulary scores.

9. IQ 60, 4 letter one syllable words -79, self-reference plural 30, other reference -31, the 31, it -37, that -59, similarities 75.

This factor seems to describe a brighter group who use less one syllable 4 letter words, use more self-reference, less references to others, use the more, it and that less.

While these factors shall not be interpreted at this point, they may be of diagnostic significance, particularly if these relations are maintained in future studies. Finally, we shall present a listing of the most frequent words used by our 2 age groups.

Table 6

Most Frequent Words Used by 6th & 10th Grades

6th Grade		Rate		10th Grade		Rate
	and	2099	4.89	I	1826	4.20
	the	1436	3.23	you	1592	3.66
	I	1356	3.15	and	1119	2.80
42,924	it	1190	2.77	43,406	the	1102	2.53
Total	you	1079	2.51	Total	not	1032	2.37
	to	1039	2.42		to	1024	2.35
	a	1017	2.31		is	978	2.25
	is	797	1.85		that	861	1.98
	that	769	1.79		a	837	1.92
	we	693	1.61		do	647	.86

We note that only two of these most frequent words are not present in both samples, the words: "it" and "we" make the first ten with the 6th grades, the words "not" and "do" make the first ten with the 10th grades.

Where does all this data take us? We found with our sampling of some 80,000 words of spoken language of 30 boys of two grades and equated intelligence, that Zipf's laws seem to be applicable to spoken as well as to written language. We found that the English language is surprisingly consistent as spoken by these children. They use about 20% different words

in these 2700 word samples. We found a number of differences in the spoken language pattern of our two age groups, such as the faster speech, the larger number of "I's," "they's," "yes's," "no's," and "how's," in the older group and "we's" in the younger group. This could be considered as base-line data in spoken language of children and such data once confirmed, may very well be used in many ways: such as in assessing individuals' deficits through psycholinguistic profiles, to help in building reading material which can be easily understood, to understand the developmental sequences in language development, to obtain national samples and last but not least, to compare various cultures with each other in their psycholinguistic development.

We are presently preparing samples of spoken language of retarded children, gifted children, school dropouts, and of the parents of our children. We want to understand the meaning of psycholinguistic indicators and learn about language development. We also are most interested in stimulating psycholinguistic studies of children in other cultures to obtain comparative data. Language after all, is the basic tool of communication among men and its usage should reveal significant information of the culture they live in.

References

Bernstein, B. Linguistic codes, hesitation phenomena and intelligence. Lang. Speech, 1962, 5, 31-46.

Brown, R, and Bellugi, Ursula. Three processes in the child's acquisition of syntax. In Eric H. Lenneberg (Ed.) New directions in the study of language. Cambridge: M.I.T. Press, 1964, Pp. 131-161.

Busemann, A. Die sprache der jugend als ausdruck der entwichlungsrhytmik. Jena: Fisher, 1925. Cited in Osgood, C. E. Method and theory in experimental psychology. New York: Oxford Univer. Press, 1953, P. 719.

Chomsky, Noam. Review of Verbal Behavior, by B. F. Skinner. Language, 1959, 35, 26-58.

Chotlos, J. W. Studies in language behavior: IV. A statistical and comparative analysis of individual written language samples. Psychol. Monogr., 1944, 56, 75-111.

Darley, F. L. Sherman, Dorothy, & Siegel, G. M. Scaling of abstraction level of single words. J. speech hear. Res., 1959, 2, 161-167.

Drieman, G. H. J. Differences between written and spoken language. Acta. Psychol. Amsterdam, 1962, 20, 36-57.

Fairbanks, H. Studies in language behavior: II. The quantitative differentiation of samples of spoken language. Psychol. Monogr., 1944, 56, No. 2, 19-38.

Flesch, R. Measuring the level of abstraction. J. appl. Psychol., 1950, 34, 384-390.

Fraisse, P. & Breyton, Madeleine. Comparisons entre les langages oral et ecrit. (Comparison between oral and written language.) Anee psychol., 1959, 59, 61-71.

Fraprie, F. R. The twenty commonest English words...from a count of 242, 423 words of English text taken from fifteen English authors and many newspapers. World almanac and book of facts for 1950 (World - Telegram, New York), P. 543.

Gleser, Goldine C., Gottschalk, L. A., & John, W. The relationship of sex and intelligence to choice or words: A normative study of verbal behavior. J. clin. Psychol., 1959, 15, 182-191.

Horowitz, M. W., & Newman, F. B. Spoken and written expression: An experimental analysis. J. abnorm. soc. Psychol., 1964, 68, 640-647.

Hymes, Dell. Functions of speech: An evolutionary approach. In Frederick C. Gruber (Ed.) Anthropology and education, University of Pennsylvania Press, 1961, Pp. 58-83.

Lenneberg, E. H. A biological perspective of language. In Eric H. Lenneberg (Ed.) New directions in the study of language. Cambridge: M.I.T. Press, 1964, Pp. 65-88.

McCarthy, Dorothea. Language development in children. In L. Carmichael (Ed.) A manual of child psychology (2nd ed.) New York: Wiley, 1954, Pp. 492-630.

McConnon, K. The situation factor in the language responses of nursery school children. Unpublished doctoral dissertation. Univer. of Minn. Cited in Dorothea McCarthy. Language development in children. In L. Carmichael (Ed.) A manual of child psychology. (2nd ed.) New York: Wiley, 1954.

Mann, M. B. Studies in language behavior: III. The quantitative differentiation of samples of written language. Psychol. Monogr., 1944, 56, 2, 41-74.

Mechel, H. C. Research in teaching composition and literature. In N. L. Gage (Ed.) Handbook of research on teaching. Chicago: Rand McNally & Co., 1963, Pp. 996-1007.

Miller, W. R. The acquisition of formal features of language. American Journal of Orthopsychiatry, 1964, 34, 862-867.

Milner, E. A study of the relationships between reading readiness in grade one school children and patterns of parent-child interaction. Child Developm. 1951, 22, 95-112.

Minifie, F. D., Darley, F. L., & Sherman, Dorothy. Temporal reliability of seven language measures. J. speech hear. Res., 1963, 6, 139-148.

Moscovici, S., & Humbert, Claudine. Etudes sur le comportement verbal: Langage oral et langage ecrit. (Studies on verbal behavior: Oral language and written language.) Psychol. Franc., 1960, 5, 175-186.

Pool, I. de Sola (Ed.) Trends in content analysis, Urbana: Univer. of Ill. Press, 1959.

Pringle, M. L. K., & Tanner, Margaret. The effects of early deprivation on speech development: A comparative study of 4 year olds in a nursery school and in residential nurseries. Lang. speech, 1958, 1, 269-287.

Rozenzweig, M. R., & McNeill, D. Inaccuracies in the semantic count of Lorge and Thorndike. Amer. J. Psychol., 1962, 75, 316-319.

Rosenzewig, M. R., & McNeill, D. Uses of the semantic count in experimental studies of verbal behavior. Amer. J. Psychol., 1962, 75, 492-495.

Smith, Madorah E. An investigation of the development of the sentence and the extent of vocabulary in young children. Univ. Ia. Stud. Child Welf., 1926, 3, No. 5.

Soparta, S. (Ed.) Psycholinguistics: A book of readings. New York: Holt, Rinehart & Winston, 1961

Starkweather, J.A. Computest: A computer language for individual testing, instruction, and interviewing. Paper read at meetings of the American Psychological Association, Los Angeles, 1964.

Strickland, R.G. The language of elementary school children: Its relationship to the language of reading textbooks and the quality of reading of selected children. Bull. Sch. educ. Ind. U., 1962, 38, 1-31.

Thorndike, E.L., & Lorge, I. A teacher's work book of 30,000 words. New York: Bureau of Publ., N.Y. Teachers College, Columbia Univer., 1944.

Verplank, W.S. The control of the content of conversation: Reinforcement of statements of opinion. J. abnorm. soc. Beh., 1955, 51, 668-676.

Voelker, C.H. The one-thousand most frequent spoken words. Quart. J. Speech, 1942, 28, 189-197.

Zipf, G.K. The psycho-biology of language. Boston: Houghton Mifflin, 1935.

Zipf, G.K. Observations of the possible effect of mental age upon the frequency distribution of words from the viewpoint of dynamic philology. J. Psychol., 1937, 4, 239-244.